W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 02:58 pm BST

PDB ID	:	2WVH
Title	:	Structural insights into the pre-reaction state of pyruvate decarboxylase from Zymomonas mobilis
Authors	:	Pei, X.Y.; Erixon, K.; Luisi, B.F.; Leeper, F.J.
Deposited on	:	2009-10-16
Resolution	:	2.30 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

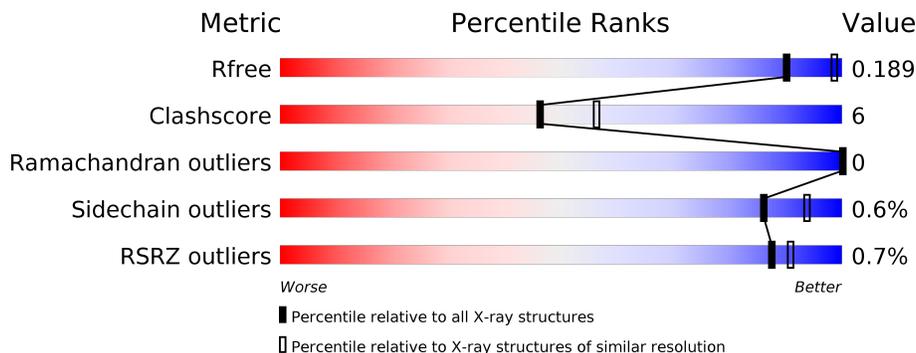|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Percentile Ranks	Value
Rfree		0.189
Clashscore		6
Ramachandran outliers		0
Sidechain outliers		0.6%
RSRZ outliers		0.7%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	568	87% 12%
1	B	568	88% 11%
1	E	568	88% 11%
1	F	568	86% 13%
1	V	568	88% 12%
1	X	568	87% 12%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	Y	568	88% 11%
1	Z	568	88% 11%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 38331 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called PYRUVATE DECARBOXYLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	566	Total 4281	C 2724	N 730	O 808	S 19	0	1	0
1	B	565	Total 4276	C 2720	N 730	O 808	S 18	0	1	0
1	E	565	Total 4273	C 2719	N 729	O 807	S 18	0	1	0
1	F	565	Total 4273	C 2719	N 729	O 807	S 18	0	1	0
1	V	565	Total 4273	C 2719	N 729	O 807	S 18	0	1	0
1	X	564	Total 4264	C 2713	N 727	O 806	S 18	0	1	0
1	Y	565	Total 4273	C 2719	N 729	O 807	S 18	0	1	0
1	Z	565	Total 4273	C 2719	N 729	O 807	S 18	0	1	0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	614	Total 614	O 614	0	0
2	B	479	Total 479	O 479	0	0
2	E	546	Total 546	O 546	0	0
2	F	522	Total 522	O 522	0	0
2	V	482	Total 482	O 482	0	0
2	X	486	Total 486	O 486	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	Y	503	Total 503	O 503	0	0
2	Z	513	Total 513	O 513	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.
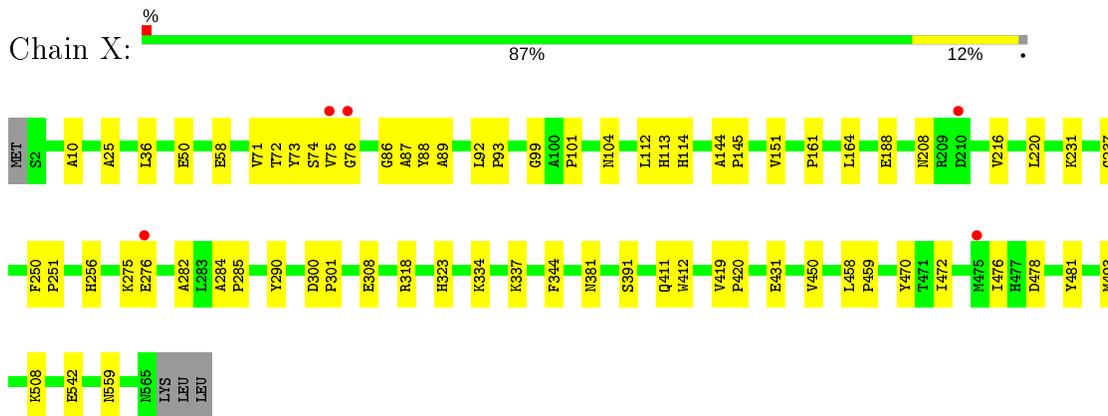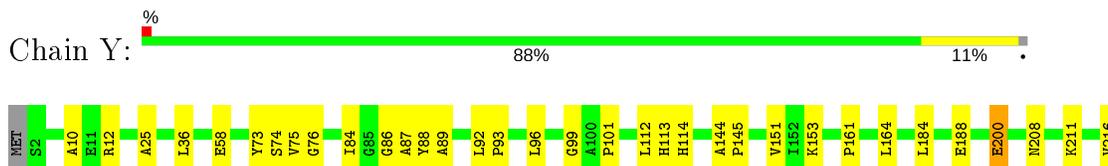
- Molecule 1: PYRUVATE DECARBOXYLASE
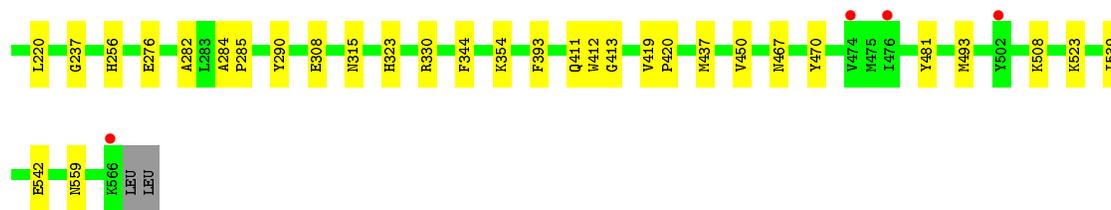

- Molecule 1: PYRUVATE DECARBOXYLASE

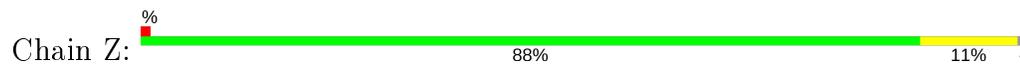- Molecule 1: PYRUVATE DECARBOXYLASE

• Molecule 1: PYRUVATE DECARBOXYLASE

Chain F:

• Molecule 1: PYRUVATE DECARBOXYLASE

Chain V:

• Molecule 1: PYRUVATE DECARBOXYLASE

Chain X:

• Molecule 1: PYRUVATE DECARBOXYLASE

Chain Y:

L220 G237 H256 E276 A282 L283 A284 P285 Y290 E308 N315 H323 R330 F344 K354 F393 Q411 W412 G413 V419 P420 M437 V450 N467 Y470 V474 M475 I476 Y481 M493 Y502 K508 K523 I539

E542 N559 K566 LEU LEU

● Molecule 1: PYRUVATE DECARBOXYLASE

Chain Z:

88%   11%

MET S2 A10 E11 R12 A25 L36 Y46 Y73 S74 V75 G76 G86 A87 Y88 A89 L92 P93 H113 H114 K118 A144 P145 V151 I152 K153 P161 L164 L184 E188 A193 F205 N208 K211 G237 F250 P251 N254

P255 H256 E276 A282 L283 A284 P285 D289 Y290 P299 E308 H323 F344 E352 L377 T388 Q411 W412 G413 H414 I415 V419 P420 L445 V450 Y470 V474 Y481 M493 Y502 K508 R541 E542 V555 N559 N565

K566 LEU LEU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	70.90Å 111.17Å 167.12Å 89.89° 90.87° 101.27°	Depositor
Resolution (Å)	16.97 – 2.30 47.66 – 2.30	Depositor EDS
% Data completeness (in resolution range)	95.6 (16.97-2.30) 93.9 (47.66-2.30)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	3.08 (at 2.29Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.171 , 0.194 0.168 , 0.189	Depositor DCC
R_{free} test set	10948 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	22.8	Xtriage
Anisotropy	0.424	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 22.0	EDS
L-test for twinning[2]	$<|L|> = 0.48$, $<L^2> = 0.31$	Xtriage
Estimated twinning fraction	0.337 for -h,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	38331	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.29% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $<|L|>$, $<L^2>$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$>5	RMSZ	#$|Z|$>5
1	A	0.38	0/4377	0.46	0/5950
1	B	0.36	1/4369 (0.0%)	0.48	1/5940 (0.0%)
1	E	0.38	0/4369	0.46	0/5940
1	F	0.41	0/4369	0.48	0/5940
1	V	0.34	0/4369	0.46	0/5940
1	X	0.34	0/4360	0.46	0/5929
1	Y	0.35	0/4369	0.46	0/5940
1	Z	0.36	0/4369	0.46	0/5940
All	All	0.36	1/34951 (0.0%)	0.47	1/47519 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	200	GLU	CB-CG	-5.75	1.41	1.52

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	565	ASN	O-C-N	5.26	131.12	122.70

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4281	0	4255	58	0
1	B	4276	0	4242	56	0
1	E	4273	0	4243	53	0
1	F	4273	0	4243	68	0
1	V	4273	0	4243	62	0
1	X	4264	0	4230	69	0
1	Y	4273	0	4243	53	0
1	Z	4273	0	4243	57	0
2	A	614	0	0	13	0
2	B	479	0	0	13	0
2	E	546	0	0	12	0
2	F	522	0	0	25	0
2	V	482	0	0	15	0
2	X	486	0	0	11	0
2	Y	503	0	0	12	0
2	Z	513	0	0	12	0
All	All	38331	0	33942	429	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

The worst 5 of 429 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:X:72:THR:CG2	1:X:75:VAL:HG23	1.69	1.23
1:X:72:THR:HG23	1:X:75:VAL:CG2	1.77	1.14
1:X:72:THR:CG2	1:X:75:VAL:CG2	2.26	1.12
1:F:308:GLU:HG3	2:F:2343:HOH:O	1.52	1.07
1:F:380:PRO:HB2	1:X:431:GLU:HB3	1.43	1.00

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	565/568 (100%)	553 (98%)	12 (2%)	0	100	100
1	B	564/568 (99%)	554 (98%)	10 (2%)	0	100	100
1	E	564/568 (99%)	554 (98%)	10 (2%)	0	100	100
1	F	564/568 (99%)	553 (98%)	11 (2%)	0	100	100
1	V	564/568 (99%)	555 (98%)	9 (2%)	0	100	100
1	X	563/568 (99%)	553 (98%)	10 (2%)	0	100	100
1	Y	564/568 (99%)	554 (98%)	10 (2%)	0	100	100
1	Z	564/568 (99%)	555 (98%)	9 (2%)	0	100	100
All	All	4512/4544 (99%)	4431 (98%)	81 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	436/437 (100%)	435 (100%)	1 (0%)	93	97
1	B	435/437 (100%)	430 (99%)	5 (1%)	73	86
1	E	435/437 (100%)	433 (100%)	2 (0%)	88	95
1	F	435/437 (100%)	431 (99%)	4 (1%)	78	89
1	V	435/437 (100%)	434 (100%)	1 (0%)	93	97
1	X	434/437 (99%)	432 (100%)	2 (0%)	88	95
1	Y	435/437 (100%)	432 (99%)	3 (1%)	84	92
1	Z	435/437 (100%)	433 (100%)	2 (0%)	88	95
All	All	3480/3496 (100%)	3460 (99%)	20 (1%)	86	94

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	308	GLU
1	F	541	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	Y	200	GLU
1	E	474	VAL
1	F	58	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 46 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	256	HIS
1	V	208	ASN
1	Z	208	ASN
1	F	323	HIS
1	F	565	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	566/568 (99%)	-0.24	6 (1%)	80	85	15, 21, 39, 65	1 (0%)
1	B	565/568 (99%)	-0.15	4 (0%)	87	91	14, 27, 49, 73	1 (0%)
1	E	565/568 (99%)	-0.23	2 (0%)	92	95	15, 23, 41, 73	1 (0%)
1	F	565/568 (99%)	-0.24	2 (0%)	92	95	15, 22, 42, 71	1 (0%)
1	V	565/568 (99%)	-0.17	6 (1%)	80	85	15, 27, 48, 66	1 (0%)
1	X	564/568 (99%)	-0.16	5 (0%)	84	88	15, 27, 48, 67	1 (0%)
1	Y	565/568 (99%)	-0.13	4 (0%)	87	91	16, 27, 45, 67	1 (0%)
1	Z	565/568 (99%)	-0.18	4 (0%)	87	91	16, 26, 45, 63	1 (0%)
All	All	4520/4544 (99%)	-0.19	33 (0%)	87	91	14, 25, 47, 73	8 (0%)

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	502	TYR	6.6
1	A	1	MET	4.2
1	X	76	GLY	4.1
1	X	75	VAL	4.0
1	Z	502	TYR	3.7

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.